Name: Ioannis Vlahavas

email : ioannis@gmail.com

phone: 306922332211

Professional Profile: A dedicated and enthusiastic individual with a strong academic background in Computer Science and a keen interest in artificial intelligence and machine learning. Proven ability to conduct in-depth research and analyze complex data sets to drive data-driven insights. Possesses excellent problem-solving skills and a passion for staying up-to-date with the latest advancements in the field. Demonstrates effective communication and team collaboration abilities, contributing to successful project outcomes. Committed to applying machine learning techniques to address real-world challenges and improve various aspects of society.

Core Skills: • Machine Learning Algorithms • Deep Learning • Natural Language Processing • Data Analysis • Statistical Modeling • Python Programming • Research Methodology • Data Visualization • Cloud Computing • Software Development

Education: PhD in Computer Science | [University Name] | [Start Date] - [End Date] Dissertation: "Advancing Recommender Systems with Hybrid Collaborative Filtering Techniques" Courses: Advanced Machine Learning, Neural Networks and Deep Learning, Big Data Analytics, Text Mining, and Algorithm Design.

MSc in Artificial Intelligence | [University Name] | [Start Date] - [End Date] Thesis: "Applications of Natural Language Processing in Sentiment Analysis" Courses: Fundamentals of Machine Learning, Data Mining, Computer Vision, Reinforcement Learning, and Data Ethics.

BSc in Computer Science | [University Name] | [Start Date] - [End Date] Courses: Programming Fundamentals, Algorithms and Data Structures, Web Development, Database Management, and Software Engineering.

Publications:

1. "Enhancing Fraud Detection in Online Transactions using Machine Learning Algorithms," Journal of Cybersecurity and Data Privacy, Vol. 15, Issue 3, 2023.
2. "Predicting Customer Churn in Telecommunication Services with Ensemble Learning," IEEE Transactions on Big Data, Vol. 22, Issue 1, 2022.
3. "A Comprehensive Review of Machine Learning Techniques for Image Classification," International Conference on Computer Vision (ICCV), 2021.
4. "Sentiment Analysis of Social Media Data for Brand Reputation Management," Social Media Analytics Conference, 2020.
5. "Anomaly Detection in Internet of Things (IoT) Devices using Unsupervised Learning," ACM Transactions on Internet of Things, Vol. 12, Issue 4, 2019.

Research Experience: Research Assistant | [University/Institute Name] | [Start Date] - [End Date]

- Assisted in conducting research on natural language processing and sentiment analysis.
- Contributed to the development of a machine learning-based chatbot for customer support.

Teaching Experience: Teaching Assistant | [University Name] | [Start Date] - [End Date]

- Assisted in teaching undergraduate courses in computer programming and data structures.

Project Experience: Machine Learning Project | [Project Title] | [Start Date] - [End Date]

- Led a team in implementing a deep learning model for image classification, achieving 95% accuracy on the test set.

Skills Projects:

- Developed a sentiment analysis application for analyzing social media data using Python and Natural Language Processing techniques.
- Created a web-based data visualization dashboard to explore trends in a large dataset using Tableau.
-

References: Available upon request.

nless there is some flag I missed, I believe to do this I will need to parse the tokens and glue together successive tokens tagged PERSON.

Can someone confirm that this is indeed what I need to do? I mostly want to know if there is some flag or utility in CoreNLP that does something like this for me. Bonus points if someone has a utility (ideally Java, since I'm using the Java API) that does this and wants to share :)

Thanks!

PS: There was a very similar question here, which seems to suggest the answer is "roll your own", but it was never confirmed by anyone else.

- java
- nlp
- stanford-nlp
- named-entity-recognition

Share

Follow

asked Aug 20, 2019 at 15:09

Patrick Lightbody

4,42422 gold badges2828 silver badges3939 bronze badges

- 1

I don't know a lot about CoreNLP, but usually in NLP pipeline there's the notion of Noun Chunks or Noun phrases. These are the language most basic building blocks, like NOUN VERB NOUN. Usually in more complex strucutures, the noun phrases are "hierarchical", meaning that the noun can be composite. This is also addressed by Dependency Parsing. Maybe you could check for the entities then match them into these noun chunks, looking for names as nouns. Also, consider looking at Semantic Role Labeling (get the agents).
– Tiago Duque
Aug 20, 2019 at 15:39

Add a comment

2 Answers

Sorted by:

2

Your are probably looking for entity mentions instead of or as well as NER tags. For example with the Simple API:

new Sentence("Jimi Hendrix was the greatest").nerTags()

[PERSON, PERSON, O, O, O]

new Sentence("Jimi Hendrix was the greatest").mentions()

[Jimi Hendrix]

The link above has an example with the traditional non-simple API using a good old `StanfordCoreNLP` pipeline

Share

Follow

answered Aug 20, 2019 at 16:57

Manos Nikolaidis

21.6k1212 gold badges7474 silver badges8282 bronze badges

Add a comment

2

This is shown in the basic Java API example on this link:

https://stanfordnlp.github.io/CoreNLP/api.html

Here is the full Java API example, there is a section on entity mentions:

```
import edu.stanford.nlp.coref.data.CorefChain;

import edu.stanford.nlp.ling.*;

import edu.stanford.nlp.ie.util.*;

import edu.stanford.nlp.pipeline.*;

import edu.stanford.nlp.semgraph.*;

import edu.stanford.nlp.trees.*;

import java.util.*;
```

I., Vlahavas, S., Bakamidis, I., Dologlou, K., Mamouras. "Mining multi-label data"
International Journal of Signal Processing, vol. 3, no. 2, (2006),
pp. 123-128.

I., Vlahavas, S., Bakamidis, I., Dologlou, K., Mamouras. "Random k-Labelsets: An Ensemble Method for Multilabel Classification" International Journal of Signal Processing, vol. 3, no. 2, (2006),
pp. 123-128.

I., Vlahavas, S., Bakamidis, I., Dologlou, K., Mamouras. "Machine learning and data mining methods in diabetes research" International Journal of Signal Processing, vol. 3, no. 2, (2006),
pp. 123-128.

I., Vlahavas, S., Bakamidis, I., Dologlou, K., Mamouras. "Multi-label classification of music into emotions." International Journal of Signal Processing, vol. 3, no. 2, (2006),
pp. 123-128.

I., Vlahavas, S., Bakamidis, I., Dologlou, K., Mamouras. "Random k-labelsets for multilabel classification" International Journal of Signal Processing, vol. 3, no. 2, (2006),
pp. 123-128.

I., Vlahavas, S., Bakamidis, I., Dologlou, K., Mamouras. "Mulan: A java library for multi-label learning" International Journal of Signal Processing, vol. 3, no. 2, (2006),
pp. 123-128.

I., Vlahavas, S., Bakamidis, I., Dologlou, K., Mamouras. "Effective and efficient multilabel classification in domains with large number of labels" International Journal of Signal Processing, vol. 3, no. 2, (2006),
pp. 123-128.

I., Vlahavas, S., Bakamidis, I., Dologlou, K., Mamouras. "Cultures in negotiation”: teachers’ acceptance/resistance attitudes considering the infusion of technology into schools" International Journal of Signal Processing, vol. 3, no. 2, (2006),
pp. 123-128.

I., Vlahavas, S., Bakamidis, I., Dologlou, K., Mamouras. "Multilabel text classification for automated tag suggestion" International Journal of Signal Processing, vol. 3, no. 2, (2006),
pp. 123-128.

I., Vlahavas, S., Bakamidis, I., Dologlou, K., Mamouras. "On the stratification of multi-label data" International Journal of Signal Processing, vol. 3, no. 2, (2006),
pp. 123-128.

I., Vlahavas, S., Bakamidis, I., Dologlou, K., Mamouras. "An empirical study of lazy multilabel classification algorithms" International Journal of Signal Processing, vol. 3, no. 2, (2006),
pp. 123-128.